Hexagonal keyboard: QWERTY vs. Typewise vs. Custom

Leo Domitrović

University of Rijeka

Abstract

Keyboard layouts with key shaped as hexagons made in this project are alternatives to standard QWERTY layout with rectangular keys widely used while typing on mobile phones. Idea was to design hexagonal keyboard layout while using standard QWERTY layout, existent Typewise layout and custom designed layout which should theoretically allow better typing speeds. All three mentioned layouts are than put to the test together with standard rectangular QWERTY layout (later Default layout).

I. INTRODUCTION

Aim of this project was to test text typing efficiency of mobile keyboard where buttons have hexagonal shape. Three different hexagonal layouts were made: first is layout similar to QWERTY, second is layout same as existent Typewise layout and third is custom designed layout. Custom layout had to be designed to have theoretical typing speed greater than previous two mentioned layouts. After implementation, text typing efficiency of all three layouts were tested and compared with default layout. Aside from text typing efficiency, usability and NASA-TLX factors were measured whereby that data were derived from questionnaires participants filled out.

II. HEXAGONAL KEYBOARD LAYOUTS DESIGNS

a. QWERTY

This keyboard layout has letter order same as the default keyboard layout and because of that it is called QWERTY. It contains lowercase letter buttons and control keys (backspace button, space button, ENTER button and Settings button). Layout is shown on the figure below. Since QWERTY letter order is often used it wouldn't be great to make any adjustments here based on dominant hand because users are so used to QWERTY order that it would require their adjustment to modified layout. Space button is placed in the middle of the screen to be in the same position for left-handed and right-handed users since whole layout can't be adjusted to them as explained earlier. ENTER button and Settings button are placed around space button. It is true that ENTER button is harder to reach for left-handed users but difference is not that big. Backspace button is in the same position as in default QWERTY layout. All buttons have the same size based on screen on which it is used.

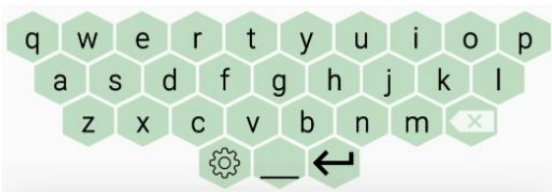


Figure 1: QWERTY Hexagonal layout

b. Typewise

Typewise layout is taken from existent Typewise layout [3]. This layout is made to optimize two-handed typing but here it is used to test its efficiency in one-handed usage. Since letter order has similarities with QWERTY letter order it didn't allow letter order to be adjusted to the dominant hand for the same reason explained for Hexagonal QWERTY layout. As it is shown below there are two buttons without any functionality and it allowed this layout to be a little adjusted to dominant hand because now space and backspace buttons can be placed on the left side of the screen for the left-handed users and the opposite for right-handed.

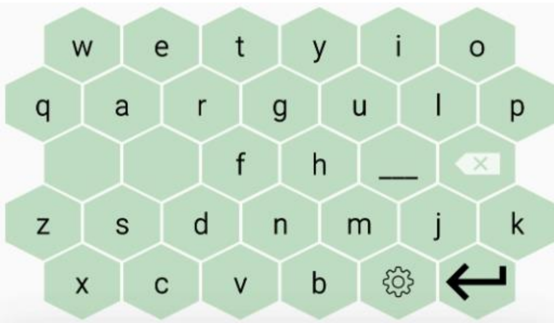


Figure 2: Typewise Hexagonal layout for right-handed users

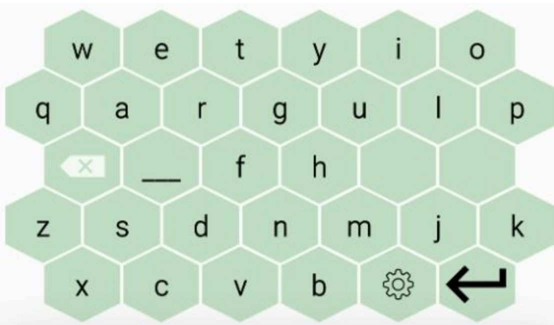


Figure 3: Typewise Hexagonal layout for left-handed users

c. Custom

This layout is designed to have the best theoretical typing speed (WPM_max) in English language between all three layouts designed in this project. Because this layout is completely unknown to users it can be fully adjusted based on dominant hand. Adjustment is made mirroring the whole layout as shown on figure below In that case, both left-handed and right handed users can have the same theoretical WPM_max and mirroring doesn't cause any difference in that matter.

Layout is designed to minimize the thumb travel. Based on digraph probabilities in English language [reference] letters that appear in more often digraph are placed closer to each other and letters in digraphs that appear rarly are placed in further positions. That is how the best WPM_max is achieved.

By studying digraph probabilities it can be seen that space appears in most of the most common digraphs and that's why space is put so close to the dominant hand. All the letters that appear in digraphs with space are placed around space or very close to it.

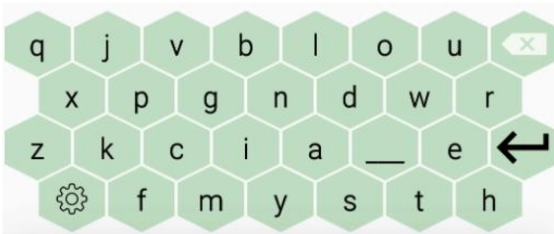


Figure 4: Custom Hexagonal layout for right-handed users

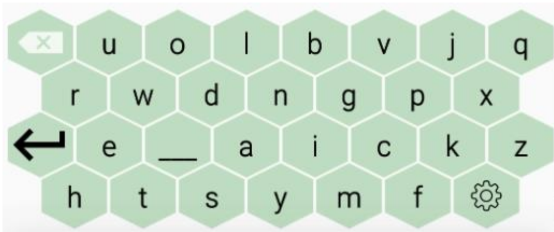


Figure 5: Custom Hexagonal layout for left-handed users

III. PARTICIPANTS

There were 9 people that participated in an experiment. Two of them are females and seven of them are male with an average of 25 years. All of them use mobile phones several times a day and have a lot of typing tasks per day. They are all non-expert users. All of the users have been right-handed. 8 participants were male and 1 was female. They all normally use Standard QWERTZ layout and do that with two hands. The study lasted around 45 minutes for each participant.

IV. APPARATUS

Participants conducted a research on different mobile devices with different versions of Android OS and different screen resolutions. Experiment was handled in two Android application called FittsTouch and Text Input Logger.

Text Input Logger uses 500 predefined phrases from English language and gives us the option of defining our custom phrases. While we do an experiment app logs information about each written phrases such as timestamp, user info, test session info, text entry metrics, task execution time, text entry metrics (calculated) (WPM, TER, CER, NCER, CPS, etc.), etc.

FittsTouch is used for ISO Multi-directional tapping test. In that test participants selected circular targets as fast and as accurate as they could. Data collected with that app is than examined and Fitts' coefficients were extracted. They are needed for predictive modeling explained later.

Before doing an experiment participants filled out a form with their personal data (name, mobile device, number of typing tasks, etc.). After a short training on each of the Hexagonal keyboard layouts exoeriment could begin. Each participant than entered 20 phrases with default device keyboard using QWERTY layout and each Hexagonal keyboard layout. They typed phrases as fast as possible, as accurate as possible while errors were allowed in submitted text. After the experiment they gave their personal opinions about keyboard layouts design, their usability, NASA-TLX factors and what Hexagonal keyboard layout they would continue to use.

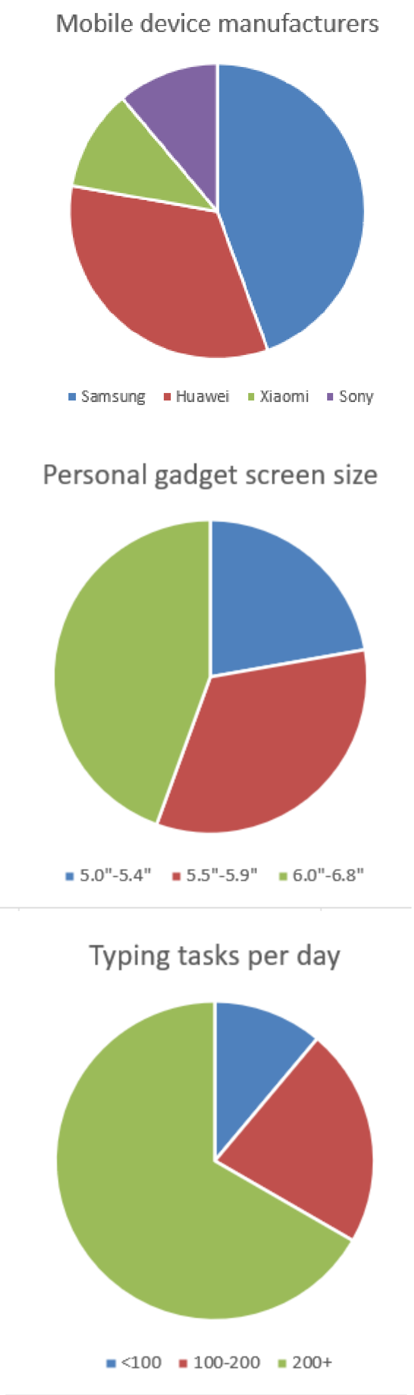


Figure 6: Number of typing tasks per day

V. PREDICTIVE MODELING

To calculate WPM_max predictive modeling with Fitts' law was used. For that we need Movement time model shown below:

$$MT = a + b \cdot \log_2(\frac{A_{ij}}{W_j} + 1)$$

Language model is also needed for predictive model.

Language model:

- Target language: English
- Character set: C = {a, …, z, space}
- Possible digraphs 27 x 27 = 729
- Digraph probabilities

First														Second Letter													
Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Space	Total
A	2	144	308	382	1	67	138	9	322	7	146	664	177	1576	1	100	-	802	683	785	87	233	57	14	319	12	50	7086
B	136	14	-	-	415	-	-	-	78	18	-	98	1	-	240	-	-	88	15	7	256	1	1	-	13	-	36	1417
C	368	-	13	-	285	-	-	412	67	-	178	108	-	1	298	-	1	71	7	154	34	-	-	-	9	-	47	2053
D	106	1	-	37	375	3	19	-	148	1	-	22	1	2	137	-	-	83	95	3	52	5	2	-	51	-	2627	3770
E	670	8	181	767	470	103	46	15	127	1	35	332	187	799	44	90	9	1314	630	316	8	172	106	87	189	2	4904	11612
F	145	-	-	-	154	86	-	-	205	-	-	69	3	-	429	-	-	188	4	102	62	-	-	-	4	-	110	1561
G	94	1	-	-	209	-	19	288	96	-	-	55	1	31	135	-	-	98	42	6	57	-	1	-	2	-	686	1901
H	1164	-	-	-	3155	-	-	1	824	-	-	5	1	-	487	2	-	91	8	165	75	-	8	-	32	-	715	6733
I	23	7	304	260	189	56	233	-	1	-	86	324	255	1110	88	42	2	272	484	558	5	165	-	15	-	18	4	4501
J	2	-	-	-	31	-	-	-	9	-	-	-	-	-	41	-	-	-	-	-	56	-	-	-	-	-	-	139
K	2	-	-	-	337	-	-	-	127	-	-	10	1	82	3	1	-	-	50	-	3	-	-	-	8	-	309	933
L	332	4	6	289	591	59	7	-	390	-	38	546	30	1	344	34	-	11	121	74	81	17	19	-	276	-	630	3900
M	394	50	-	-	530	6	-	-	165	-	-	4	28	4	289	77	-	-	53	2	85	-	-	-	19	-	454	2160
N	100	2	98	1213	512	5	771	5	135	8	63	80	-	54	349	-	3	2	148	378	49	3	2	2	115	-	1152	5249
O	65	67	61	119	34	80	9	1	88	3	123	218	417	598	336	138	-	812	195	415	1115	136	398	2	47	5	294	5776
P	142	-	1	-	280	1	-	24	97	-	-	169	-	-	149	64	-	110	48	40	68	-	3	-	14	-	127	1337
Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	66	-	-	-	-	-	-	66
R	289	10	22	133	1139	13	59	21	309	-	53	71	65	106	504	9	-	69	318	190	89	22	5	-	145	-	1483	5124
S	196	9	47	-	626	-	1	328	214	-	57	48	31	16	213	107	8	-	168	754	175	-	32	-	34	-	2228	5292
T	259	2	31	1	583	1	2	3774	252	-	-	75	1	2	331	-	-	187	209	154	132	-	84	-	121	1	2343	8545
U	45	53	114	48	71	10	148	-	65	-	-	247	87	278	3	49	1	402	299	492	-	-	-	1	7	3	255	2678
V	27	-	-	-	683	-	-	-	109	-	-	-	-	-	33	-	-	-	-	-	1	-	-	-	11	-	-	864
W	595	3	-	6	285	-	-	472	374	-	1	12	-	103	264	-	-	35	21	4	2	-	-	-	-	-	326	2503
X	17	-	9	-	9	-	-	-	10	-	-	-	-	-	1	22	-	-	-	23	8	-	-	-	-	-	21	120
Y	11	10	-	-	152	-	1	1	32	-	-	7	1	-	339	16	-	-	81	2	1	-	2	-	-	-	1171	1827
Z	3	-	-	-	26	-	-	-	2	-	-	4	-	-	2	-	-	-	3	-	-	-	-	-	3	9	2	54
Space	1882	1033	864	515	423	1059	453	1388	237	93	152	717	876	478	721	588	42	494	1596	3912	134	116	1787	-	436	2	-	19998
Total	7069	1418	2059	3770	11645	1549	1906	6739	4483	131	932	3885	2163	5241	5781	1339	66	5129	5278	8536	2701	870	2507	121	1855	52	19974	107199

Figure 6: Digraph count table [1]

From this table probabilities were extracted by dividing digraph count with total number of letters.

Next, pairing of movement time model and language time model was done with formulas below:

- Average thumb movement time between two buttons on a keyboard

$$CT = \sum_{i \in C} \sum_{j \in C} (P_{ij} \cdot MT_{ij})$$

where P_{ij} is probability of specific digraph, MT_{ij} is movement time between buttons, CT stands for *Character Time*

- Characters per second

$$CPS_{max} = \frac{1}{CT} \ [s^{-1}]$$

- Predicted typing efficiency (WPM)

$$WPM_{max} = CPS_{max} \cdot \frac{60}{5}$$

VI. PERFORMANCE ESTIMATION OF HEXAGONAL KEYBOARD LAYOUTS

For all hexagonal keyboard layouts theoretical typing speeds (WPM_max) were calculated. Firstly, it was done for QWERTY and Typewise keyboard layouts. Based on that results Custom layout was designed to have best WPM_max value of other two hexagonal layouts. Actual button locations and width on mobile screen were used in modeling. In that way, when I calculate distance of two targets (keyboard buttons) that are actual distance on screen. After rendering keyboard on Samsung Galaxy S5 and finding out actual button locations on screen theoretical WPM_{max} could be calculated.

For testing, calculation was done with Fitts' coefficients from literature and with Fitts' coefficients extracted from FittsTouch experiment.

From literature [4]: a = 0.1154, b = 0.1098

From experiment: a = 0.3887, b = 0.0639

a. Performance of Hexagonal QWERTY layout

With coefficients from literature theoretical speed of 36.485 WPM was calculated. But with coefficients from experiment speed is 23.39 WPM In the next subsection it can be seen how it compares to the other two layout.

b. Performance of Hexagonal Typewise layout

Here we will have two WPM_{max} values because layout has an adjustment based on dominant hand as explained earlier.

For right-handed users theoretical WPM_{max} value is 39.314 with coefficients from literature and 24.447 with coefficients from experiment. From that we can conclude that Hexagonal Typewise layout is better than Hexagonal QWERTY layout even though existent Typewise layout is adjusted for two-thumb typing.

Left-handed theoretical typing speed was calculated to be 41.256 WPM with coefficients from literature literature and 24.037 with coefficients from experiment which is even better than Hexagonal QWERTY layout and Hexagonal Typewise layout for right-handed users.

c. Performance of Hexagonal Custom layout

Theoretical typing efficiency of this layout was calculated to be 44.274 WPM with coefficients from literature and 25.036 with coefficients from experiment.. Since the idea of this layout was to have the best WPM_{max} value of all hexagonal layouts designed in this project it can be concluded that it is successfully achieved. As it was explained earlier, left-handed and right-handed users should have the same WPM_{max} value and that is the case. It can be said that left-handed and right-handed users have the same theoretical typing speed and there are no differences in that matter like it was with Hexagonal Typewise layout.

VII. RESULTS

a. WPM

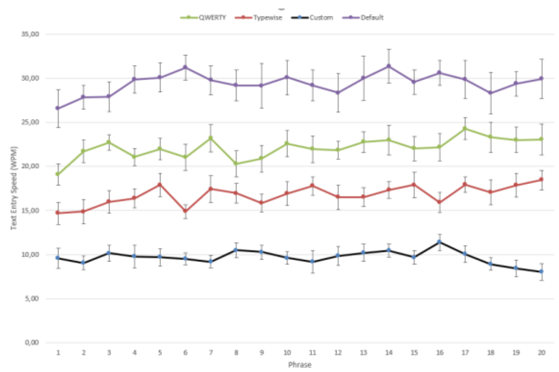


Figure 7: WPM by phrase

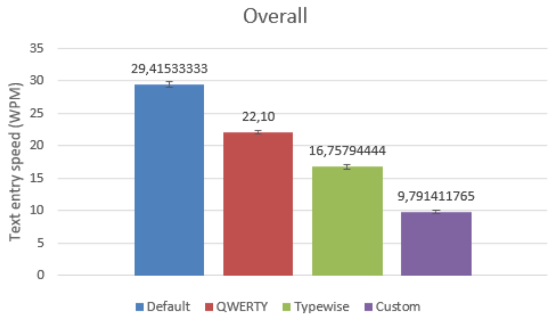


Figure 8: Overall WPM

Graphs above shows an average WPM values by phrase and overall for each keyboard layout tests. For each layout there isn't a trend od improving WPM after each phrase but there are noticeable differences in WPM if we compare them to the other layouts. Default layout has the best WPM across all phrases. Hexagonal QWERTY layout is the first behind because of the familiar letter order by since it is smaller than default users needed to adjust to it. Third best layout here Hexagonal Typewise layout and that is due to the new letter order comparing to QWERTY. It is similar to QWERTY but has some differences. Custom layout was the worst in text typing speed because of completely new letter order. Users would need a very long time to adjust to it and amount of training they did before the experiment wasn't enough.

To test if there are any statistical differences in WPM values One-way ANOVA with repeated measures was used.

Estimates

Measure: WPM

Speed	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	29,415	,431	28,564	30,267
2	22,099	,326	21,456	22,743
3	16,758	,284	16,198	17,318
4	9,640	,217	9,213	10,068

Figure 9: Descriptive statistics

1- Default
2- QWERTY
3- Typewise
4- Custom

Tests of Within-Subjects Effects

Measure: WPM

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Speed	Sphericity Assumed	37765,208	3	12588,403	1040,347	<,001	,853
	Greenhouse-Geisser	37765,208	2,396	15760,574	1040,347	<,001	,853
	Huynh-Feldt	37765,208	2,431	15535,245	1040,347	<,001	,853
	Lower-bound	37765,208	1,000	37765,208	1040,347	<,001	,853
Error(Speed)	Sphericity Assumed	6497,805	537	12,100			
	Greenhouse-Geisser	6497,805	428,917	15,149			
	Huynh-Feldt	6497,805	435,138	14,933			
	Lower-bound	6497,805	179,000	36,301			

Since p<0.001, which is smaller than 0.05, it can be concluded that there are significant differences between WPM values of each layout. To see where the differences occur Friedman test Is used.

Pairwise Comparisons

Measure: WPM

(I) Speed	(J) Speed	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
1	2	7,316*	,428	<,001	6,173	8,459
	3	12,657*	,426	<,001	11,522	13,793
	4	19,775*	,420	<,001	18,654	20,896
2	1	-7,316*	,428	<,001	-8,459	-6,173
	3	5,342*	,323	<,001	4,481	6,202
	4	12,459*	,307	<,001	11,639	13,280
3	1	-12,657*	,426	<,001	-13,793	-11,522
	2	-5,342*	,323	<,001	-6,202	-4,481
	4	7,118*	,259	<,001	6,427	7,809
4	1	-19,775*	,420	<,001	-20,896	-18,654
	2	-12,459*	,307	<,001	-13,280	-11,639
	3	-7,118*	,259	<,001	-7,809	-6,427

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Comparison between each of the layouts shows that there is significant difference between all of them.

b. TER

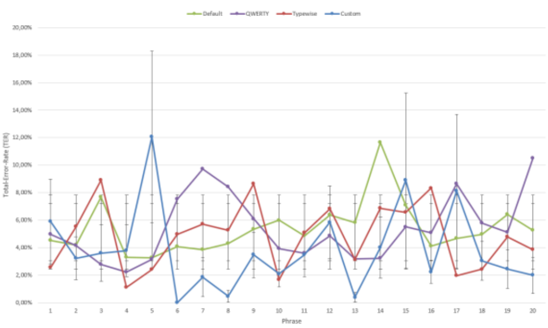


Figure 11: TER by phrase

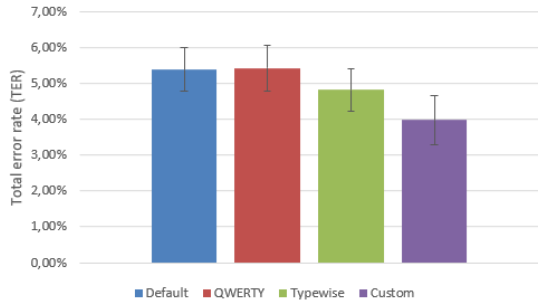


Figure 12: Overall TER

To test if there are any statistical differences in WPM values One-way ANOVA with repeated measures was used.

Estimates

Measure: TER

Errors	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
1	5,375	,607	4,176	6,573
2	5,413	,643	4,144	6,683
3	4,816	,596	3,640	5,992
4	3,966	,680	2,625	5,308

Figure 13: Descriptive statistics

Tests of Within-Subjects Effects

Measure: TER

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Errors	Sphericity Assumed	246,115	3	82,038	1,330	,264
	Greenhouse-Geisser	246,115	2,920	84,296	1,330	,264
	Huynh-Feldt	246,115	2,973	82,776	1,330	,264
	Lower-bound	246,115	1,000	246,115	1,330	,250
Error(Errors)	Sphericity Assumed	33126,137	537	61,687		
	Greenhouse-Geisser	33126,137	522,616	63,385		
	Huynh-Feldt	33126,137	532,217	62,242		
	Lower-bound	33126,137	179,000	185,062		

Figure 14: One-way ANOVA with repeated measures test

One-way ANOVA test showed that there are NO statistically significant differences between TER values of each layout. That was also proved in multiple

Pairwise Comparisons

Measure: TER

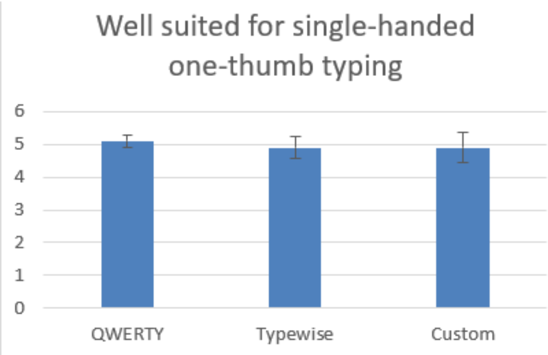
(I) Errors	(J) Errors	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
1	2	-,039	,835	1,000	-2,266	2,189
	3	,559	,749	1,000	-1,440	2,557
	4	1,408	,848	,590	-,853	3,670
2	1	,039	,835	1,000	-2,189	2,266
	3	,597	,787	1,000	-1,503	2,698
	4	1,447	,864	,575	-,859	3,753
3	1	-,559	,749	1,000	-2,557	1,440
	2	-,597	,787	1,000	-2,698	1,503
	4	,850	,877	1,000	-1,490	3,189
4	1	-1,408	,848	,590	-3,670	,853
	2	-1,447	,864	,575	-3,753	,859
	3	-,850	,877	1,000	-3,189	1,490

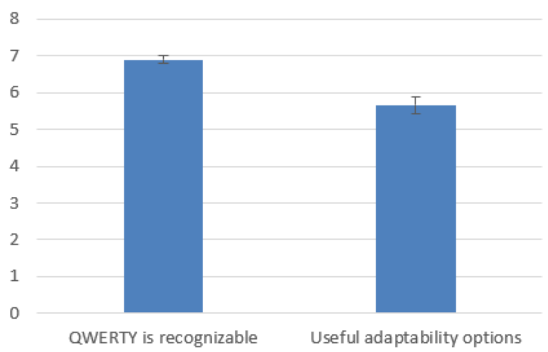
Based on estimated marginal means

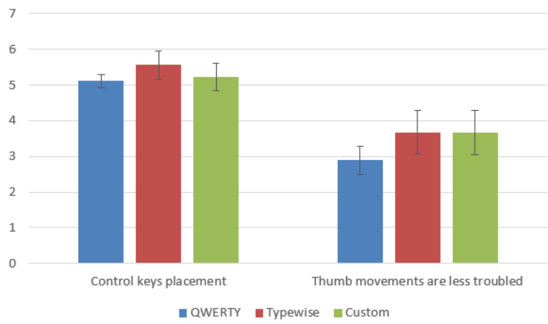
a. Adjustment for multiple comparisons: Bonferroni.

c. Design considerations

Based on subjective opinions participants gave after the experiment analysis looks like this:







Questions that were asked were:

1. Is each hexagonal layout well suited for single-handed one-thumb typing?
2. Is standard QWERTY layout recognizable from Hexagonal QWERTY layout?
3. Are adaptability options useful?
4. Are control keys placed on comfortable places?
5. Are thumb movements less troubled using Hexagonal keyboard layouts?

To see if there is any statistical difference between layouts on the first, the third and last question Friedman test is used.

Test Statistics[a]

N	9
Chi-Square	,867
df	2
Asymp. Sig.	,648

a. Friedman Test

Figure 15: Well-suited test

Test Statistics[a]

N	9
Chi-Square	4,160
df	2
Asymp. Sig.	,125

a. Friedman Test

Figure 16: Control keys test

Test Statistics[a]

N	9
Chi-Square	6,778
df	2
Asymp. Sig.	,034

a. Friedman Test

Figure 17: Movements test

- There was NO statistically significant difference in perceived well-adjustment to one-thumb typing depending on which keyboard layout was used, $\chi^2(2) = 0.867$, $p = 0.648$
- There was NO statistically significant difference in perceived control keys correct placement depending on which keyboard layout was used, $\chi^2(2) = 4.160$, $p = 0.125$
- There was a statistically significant difference in perceived comfortable thumb movement depending on which keyboard layout was used, $\chi^2(2) = 6.778$, $p < 0.034$

d. NASA-TLX factors

After the experiment users were asked to fill out questionnaire in which they gave their subjective opinions about mental demand, physical demand, frustration, performance and effort. They did that for default layout and all three hexagonal keyboard layouts.

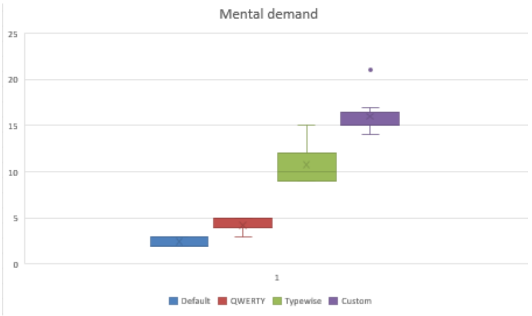


Figure 18: Mental demand comparison

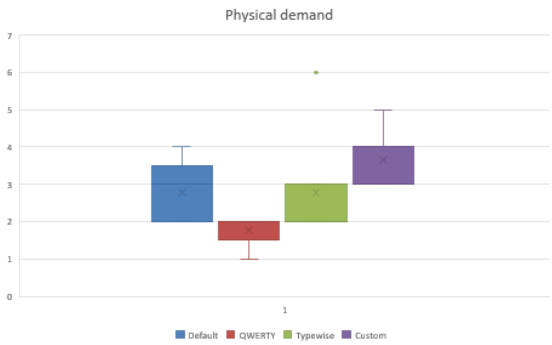


Figure 19: Physical demand comparison

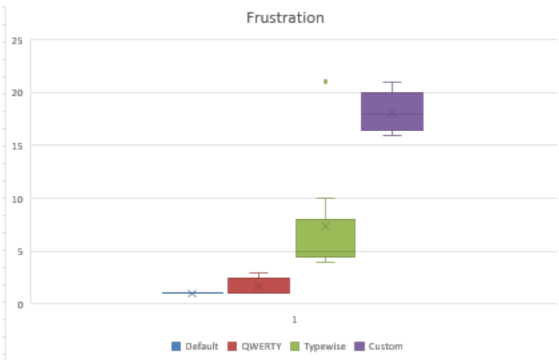


Figure 20: Frustration comparison

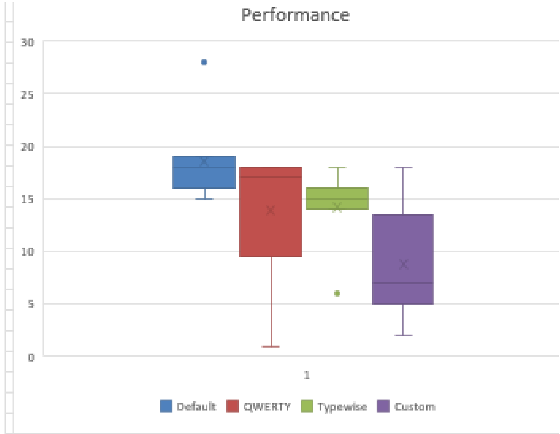


Figure 21: Performance comparison

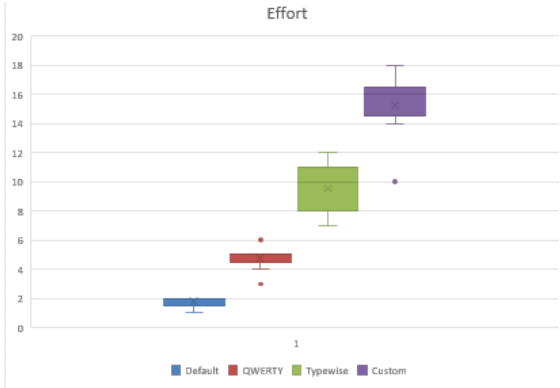


Figure 22: Effort comparison

Test Statistics[a]

N	9
Chi-Square	27,000
df	3
Asymp. Sig.	<,001

a. Friedman Test

Figure 23: Mental demand test

Test Statistics[a]

N	9
Chi-Square	9,579
df	2
Asymp. Sig.	,008

a. Friedman Test

Figure 24: Physical demand test

Test Statistics[a]

N	9
Chi-Square	26,000
df	3
Asymp. Sig.	<,001

a. Friedman Test

Figure 25: Frustration test

Test Statistics[a]

N	9
Chi-Square	10,291
df	3
Asymp. Sig.	,016

a. Friedman Test

Figure 26: Performance test

Test Statistics[a]

N	9
Chi-Square	26,730
df	3
Asymp. Sig.	<,001

a. Friedman Test

Figure 27: Effort test

To test if there is any difference between between each keyboard layout for each of the dependent variables shown on the graph I used Friedman test.

- There was a statistically significant difference in perceived mental demand depending on which keyboard layout was used, $\chi^2(2) = 27$, $p < 0.001$
- There was a statistically significant difference in perceived physical demand depending on which keyboard layout was used, $\chi^2(2) = 9.579$, $p = 0.008$
- There was a statistically significant difference in perceived frustration depending on which keyboard layout was used, $\chi^2(2) = 26$, $p < 0.001$
- There was a statistically significant difference in perceived performance depending on which keyboard layout was used, $\chi^2(2) = 10.291$, $p = 0.016$
- There was a statistically significant difference in perceived effort depending on which keyboard layout was used, $\chi^2(2) = 26.73$, $p < 0.001$

To examine where the differences actually occur, we need to run separate Wilcoxon signed-rank tests on the different combinations of related groups. In this case:

- Default vs QWERTY (on Mental demand, Physical demand, Frustration and Effort)
- Default vs Typewise (on Mental demand, Physical demand, Frustration and Effort)
- Default vs Custom (on Mental demand, Physical demand, Frustration and Effort)
- QWERTY vs Typewise (on Mental demand, Physical demand, Frustration and Effort)
- QWERTY vs Custom (on Mental demand, Physical demand, Frustration and Effort)
- Typewise vs Custom (on Mental demand, Physical demand, Frustration and Effort)

Because we are making multiple comparisons we need to use a Bonferroni adjustment on the results we get. To calculate the Bonferroni adjustment we take significance level we were initially using (0.05) and divide by the number of tests you are running. Here we made 6 tests so we divided 0.05 by 6 and that means that we have statistically significant difference if p-value is lower than 0.0083.

Test Statistics[a]

	Mental demand QWERTY - Mental demand Default	Mental demand Typewise - Mental demand Default	Mental demand Custom - Mental demand Default	Mental demand Typewise - Mental demand QWERTY	Mental demand Custom - Mental demand QWERTY	Mental demand Custom - Mental demand Typewise
Z	-2,701[b]	-2,673[b]	-2,684[b]	-2,673[b]	-2,677[b]	-2,668[b]
Asymp. Sig. (2-tailed)	,007	,008	,007	,008	,007	,008

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

Figure 28: Mental demand pairwise tests

- There was a statistically significant difference in perceived mental demand between Default and QWERTY keyboard layouts, Z = -2.701, $p = 0.007$
- There was a statistically significant difference in perceived mental demand between Default and Typewise keyboard layouts, Z = -2.673, $p = 0.008$
- There was a statistically significant difference in perceived mental demand between Default and Custom keyboard layouts, Z = -2.684, $p = 0.007$
- There was a statistically significant difference in perceived mental demand between QWERTY and Typewise keyboard layouts, Z = -2.673, $p = 0.008$
- There was a statistically significant difference in perceived mental demand between QWERTY and Custom keyboard layouts, Z = -2.677, $p = 0.007$
- There was a statistically significant difference in perceived mental demand between Typewise and Custom keyboard layouts, Z = -2.673, $p = 0.008$

Test Statistics[a]

	Phyisical demand QWERTY - Phyisical demand Default	Phyisical demand Typewise - Phyisical demand Default	Phyisical demand Custom - Phyisical demand Default	Phyisical demand Typewise - Phyisical demand QWERTY	Phyisical demand Custom - Phyisical demand QWERTY	Phyisical demand Custom - Phyisical demand Typewise
Z	-2,251[b]	,000[c]	-2,060[d]	-2,060[d]	-2,754[d]	-1,543[d]
Asymp. Sig. (2-tailed)	,024	1,000	,039	,039	,006	,123

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

c. The sum of negative ranks equals the sum of positive ranks.

d. Based on negative ranks.

Figure 29: Physical demand pairwise tests

- There was NO statistically significant difference in perceived physical demand between Default and QWERTY keyboard layouts, $Z = -2.251$, $p = 0.024$
- There was NO statistically significant difference in perceived physical demand between Default and Typewise keyboard layouts, $Z = -0.000$, $p = 1.000$
- There was NO statistically significant difference in perceived physical demand between Default and Custom keyboard layouts, $Z = -2.060$, $p = 0.039$
- There was a statistically significant difference in perceived physical demand between QWERTY and Typewise keyboard layouts, $Z = -2.060$, $p = 0.039$
- There was a statistically significant difference in perceived physical demand between QWERTY and Custom keyboard layouts, $Z = -2.754$, $p = 0.006$
- There was NO statistically significant difference in perceived physical demand between Typewise and Custom keyboard layouts, $Z = -1.543$, $p = 0.123$

Test Statistics[a]

	Frustration QWERTY - Frustration Default	Frustration Typewise - Frustration Default	Frustration Custom - Frustration Default	Frustration Typewise - Frustration QWERTY	Frustration Custom - Frustration QWERTY	Frustration Custom - Frustration Typewise
Z	-1,857[b]	-2,680[b]	-2,675[b]	-2,673[b]	-2,677[b]	-2,527[b]
Asymp. Sig. (2-tailed)	,063	,007	,007	,008	,007	,012

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

Figure 30: Frustration pairwise tests

- There was NO statistically significant difference in perceived frustration between Default and QWERTY keyboard layouts, $Z = -1.857$, $p = 0.063$
- There was a statistically significant difference in perceived frustration between Default and Typewise keyboard layouts, $Z = -2.680$, $p = 0.007$
- There was a statistically significant difference in perceived frustration between Default and Custom keyboard layouts, $Z = -2.675$, $p = 0.007$
- There was a statistically significant difference in perceived frustration between QWERTY and Typewise keyboard layouts, $Z = -2.673$, $p = 0.008$
- There was a statistically significant difference in perceived frustration between QWERTY and Custom keyboard layouts, $Z = -2.677$, $p = 0.007$
- There was NO statistically significant difference in perceived frustration between Typewise and Custom keyboard layouts, $Z = -2.527$, $p = 0.012$

Test Statistics[a]

	Effort QWERTY - Effort Default	Effort Typewise - Effort Default	Effort Custom - Effort Default	Effort Typewise - Effort QWERTY	Effort Custom - Effort QWERTY	Effort Custom - Effort Typewise
Z	-2,714[b]	-2,670[b]	-2,673[b]	-2,680[b]	-2,687[b]	-2,539[b]
Asymp. Sig. (2-tailed)	,007	,008	,008	,007	,007	,011

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

Figure 31: Effort pairwise tests

- There was a statistically significant difference in perceived effort between Default and QWERTY keyboard layouts, $Z = -2.714$, $p = 0.007$
- There was a statistically significant difference in perceived effort between Default and Typewise keyboard layouts, $Z = -2.670$, $p = 0.008$
- There was a statistically significant difference in perceived effort between Default and Custom keyboard layouts, $Z = -2.673$, $p = 0.008$
- There was a statistically significant difference in perceived effort between QWERTY and Typewise keyboard layouts, $Z = -2.680$, $p = 0.007$
- There was a statistically significant difference in perceived effort between QWERTY and Custom keyboard layouts, $Z = -2.687$, $p = 0.007$
- There was NO statistically significant difference in perceived effort between Typewise and Custom keyboard layouts, $Z = -2.539$, $p = 0.011$

Because we are making multiple comparisons we need to use a Bonferroni adjustment on the results we get. To calculate the Bonferroni adjustment we take significance level we were initially using (0.05) and divide by the number of tests you are running. Here we made 6 tests so we divided 0.05 by 6 and that means that we have statistically significant difference if p-value is lower than 0.0083.

e. Usability

After the experiment participants also filled out questionnaire in which they gave their subjective opinions about usability, learnability, satisfaction, and engagement. They did that for all three hexagonal keyboard layouts.

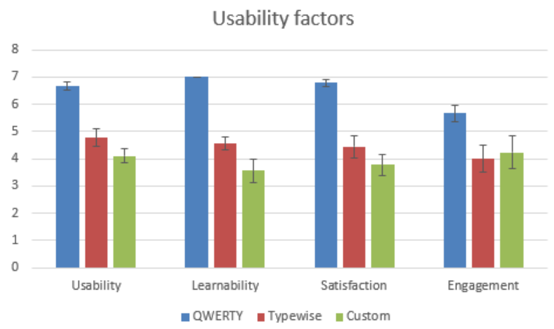


Figure 32: Usability attributes

To test if there are any statistical differences between each keyboard layout for each of the usability attributes shown on the graph I used Friedman test.

Test Statistics[a]

N	9
Chi-Square	16,750
df	2
Asymp. Sig.	<,001

a. Friedman Test

Figure 33: Usability test

Test Statistics[a]

N	9
Chi-Square	15,697
df	2
Asymp. Sig.	<,001

a. Friedman Test

Figure 34: Learnability test

Test Statistics[a]

N	9
Chi-Square	2,480
df	2
Asymp. Sig.	,289

a. Friedman Test

Figure 35: Engagement test

Test Statistics[a]

N	9
Chi-Square	15,697
df	2
Asymp. Sig.	<,001

a. Friedman Test

Figure 36: Satisfaction test

- There was a statistically significant difference in perceived usability depending on which keyboard layout was used, $\chi^2(2) = 16.750$, $p < 0.001$
- There was a statistically significant difference in perceived learnability depending on which keyboard layout was used, $\chi^2(2) = 15.697$, $p < 0.001$
- There was a statistically significant difference in perceived satisfaction depending on which keyboard layout was used, $\chi^2(2) = 15.697$, $p < 0.001$
- There was NO statistically significant difference in perceived engagement depending on which keyboard layout was used, $\chi^2(2) = 2.480$, $p = 0.289$

To examine where the differences actually occur, we need to run separate Wilcoxon signed-rank tests on the different combinations of related groups. In this case:

- QWERTY vs Typewise (on Usability, Learnability and Satisfaction)
- QWERTY vs Custom (on Usability, Learnability and Satisfaction)
- Typewise vs Custom (on Usability, Learnability and Satisfaction)

Because we are making multiple comparisons we need to use a Bonferroni adjustment on the results we get. To calculate the Bonferroni adjustment we take significance level we were initially using (0.05) and divide by the number of tests you are running. Here we made 3 tests so we divided 0.05 by 3 and that means that we have statistically significant difference if p-value is lower than 0.017.

Test Statistics[a]

	Usability Typewise - Usability QWERTY	Usability Custom - Usability QWERTY	Usability Custom - Usability Typewise
Z	-2,701[b]	-2,699[b]	-2,121[b]
Asymp. Sig. (2-tailed)	,007	,007	,034

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Figure 37: Pairwise comparison

- There was a statistically significant difference in perceived usability between QWERTY and Typewise keyboard layouts, Z = -2.701, $p = 0.007$
- There was a statistically significant difference in perceived usability between QWERTY and Custom keyboard layouts, Z = -2.699, $p = 0.007$
- There was NO statistically significant difference in perceived usability between Typewise and Custom keyboard layouts, Z = -2.121, $p = 0.034$

Test Statistics[a]

	Learnability Typewise - Learnability QWERTY	Learnability Custom - Learnability QWERTY	Learnability Custom - Learnability Typewise
Z	-2,724[b]	-2,692[b]	-1,725[b]
Asymp. Sig. (2-tailed)	,006	,007	,084

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Figure 38: Pairwise comparison

- There was a statistically significant difference in perceived learnability between QWERTY and Typewise keyboard layouts, Z = -2.724, $p = 0.006$
- There was a statistically significant difference in perceived learnability between QWERTY and Custom keyboard layouts, Z = -2.692, $p = 0.007$
- There was NO statistically significant difference in perceived learnability between Typewise and Custom keyboard layouts, Z = -1.725, $p = 0.084$

Test Statistics[a]

	Satisfaction Typewise - Satisfaction QWERTY	Satisfaction Custom - Satisfaction QWERTY	Satisfaction Custom - Satisfaction Typewise
Z	-2,687[b]	-2,694[b]	-1,667[b]
Asymp. Sig. (2-tailed)	,007	,007	,096

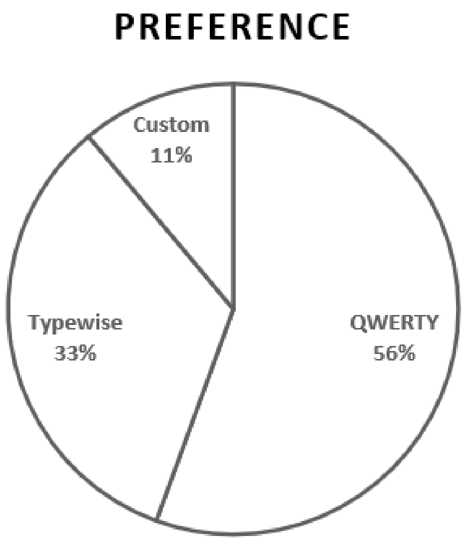
a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Figure 39: Pairwise comparison

- There was a statistically significant difference in perceived satisfaction between QWERTY and Typewise keyboard layouts, Z = -2.687, $p = 0.007$
- There was a statistically significant difference in perceived satisfaction between QWERTY and Custom keyboard layouts, Z = -2.694, $p = 0.007$
- There was NO statistically significant difference in perceived satisfaction between Typewise and Custom keyboard layouts, Z = -1.667, $p = 0.096$

f. Users preference



VIII. CONCLUSION

In this project three different hexagonal layouts were implemented where one is custom successfully designed to have the best theoretical WPM_max. They are put to the test against default QWERTY layout. Unsurprisingly, default layout was better in average WPM. It is better because every participant probably uses default layout at least 100 times a day for a very long time and neither hexagonal layout could beat that familiarity. Hexagonal QWERTY layout is right behind mostly because of the familiar letter order. Hexagonal Typewise layout is the third best in average WPM. Its letter order is similar to QWERTY but there are difference and adjustment is needed. Custom layout gave the worst results in an experiment but that’s expected because that layout was completely new to all users. A lot of typing tasks with it is required to make use of it as it was needed for default layout. It is very hard to immediately adjust to the completely new layout.

After a lot of time theoretical WPM_max could be achieved.

IX. REFERENCES

[1] Soukoreff, R. W. & MacKenzie, I. S. (1995). Theoretical upper and lower bounds on typing speed using a stylus and soft keyboard. Behaviour & Information Technology, 14, 370-379.

[2] http://www.yorku.ca/mack/FittsLawSoftware/

[3] https://www.typewise.app/

[4] Xiaojun Bi, Yang Li, and Shumin Zhai. 2013. FFitts law: modeling finger touch with fitts' law. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13). Association for Computing Machinery, New York, NY, USA, 1363–1372.